

Evaluation of 1-D tracer concentration profile in a small river by means of Multi-Layer Perceptron Neural Networks

A. Piotrowski¹, S. G. Wallis², J. J. Napiórkowski¹, and P. M. Rowiński¹

¹Institute of Geophysics, Polish Academy of Sciences, Warsaw, Poland

²Heriot-Watt University, Edinburgh, UK

Received: 24 August 2007 – Published in Hydrol. Earth Syst. Sci. Discuss.: 30 August 2007

Revised: 14 November 2007 – Accepted: 21 November 2007 – Published: 7 December 2007

Abstract. The prediction of temporal concentration profiles of a transported pollutant in a river is still a subject of ongoing research efforts worldwide. The present paper is aimed at studying the possibility of using Multi-Layer Perceptron Neural Networks to evaluate the whole concentration versus time profile at several cross-sections of a river under various flow conditions, using as little information about the river system as possible. In contrast with the earlier neural networks based work on longitudinal dispersion coefficients, this new approach relies more heavily on measurements of concentration collected during tracer tests over a range of flow conditions, but fewer hydraulic and morphological data are needed. The study is based upon 26 tracer experiments performed in a small river in Edinburgh, UK (Murray Burn) at various flow rates in a 540 m long reach. The only data used in this study were concentration measurements collected at 4 cross-sections, distances between the cross-sections and the injection site, time, as well as flow rate and water velocity, obtained according to the data measured at the 1st and 2nd cross-sections.

The four main features of concentration versus time profiles at a particular cross-section, namely the peak concentration, the arrival time of the peak at the cross-section, and the shapes of the rising and falling limbs of the profile are modeled, and for each of them a separately designed neural network was used. There was also a variant investigated in which the conservation of the injected mass was assured by adjusting the predicted peak concentration. The neural network methods were compared with the unit peak attenuation curve concept.

In general the neural networks predicted the main features of the concentration profiles satisfactorily. The predicted peak concentrations were generally better than those obtained using the unit peak attenuation method, and the

method with mass-conservation assured generally performed better than the method that did not account for mass-conservation. Predictions of peak travel time were also better using the neural networks than the unit peak attenuation method. Including more data into the neural network training set clearly improved the prediction of the shapes of the concentration profiles. Similar improvements in peak concentration were less significant and the travel time prediction appeared to be largely unaffected.

1 Introduction

For many years researchers have been interested in modeling the transport of pollutants, which, if released to a river, could endanger its ecosystem and peoples' health (Taylor, 1954; Fischer et al., 1979; Rutherford, 1994). In most cases interest is focused on forecasts of the peak concentration that would occur at particular locations, the arrival time of the peak and the duration of occurrence of dangerous pollutant levels. When adequate morphological and hydraulic data are available, one may apply the well-known Advection-Dispersion equation (Taylor, 1954) or its extension in the form of the Transient Storage Model (Thackston and Schnelle (1970); Czernuszenko and Rowiński (1997); Czernuszenko et al. (1998); Hart (1995); Manson (2000); Cheong and Seo (2003); Rowiński et al. (2003); De Smedt et al. (2005); Guymer and Dutton, 2005) to evaluate concentration versus time profiles at different distances from the pollutant release point. But such detailed measurements of, for example dispersion coefficients and travel times, are neither easy nor cheap to perform.

Recently it was shown that Multi-Layer Perceptron (MLP) Neural Networks may facilitate the evaluation of longitudinal dispersion coefficients in rivers where no prior knowledge of their mass transport characteristics was available (Kashefipour et al., 2002; Rowiński et al., 2005; Wallis et al., 2007).

Correspondence to: P. Rowiński
(pawelr@igf.edu.pl)

Table 1. Description of the data. TR – TRaining data set; VL – VaLidation data set; TE – TEsting data set.

Experiment number	Data division D9	Data division D5	Q (l/s) at CS1	U (m/s) CS1 - CS2	Tracer mass released (g)	Total number of collected measurements	Cross-sections covered
2	TR	TE	68	0.196	0.05	78	1.2
4	VL	VL	44	0.157	0.05	103	1.2
5	TR	TR	48	0.157	0.05	101	1.2
6	TE	TE	128	0.281	0.05	62	1.2
7	TE	TE	134	0.299	0.10	74	1.2
8	TR	TE	46	0.164	0.05	106	1.2.3
9	VL	VL	35	0.153	0.05	155	1.2.3
10	TE	TE	56	0.181	0.10	142	1.2.3
15	VL	TE	49	0.184	0.05	148	1.2.4
16	TR	TR	16	0.084	0.05	188	1.2.3.4
17	TR	TE	14	0.085	0.05	189	1.2.3.4
18	TE	TE	33	0.128	0.10	220	1.2.3.4
20	TR	TR	261	0.441	0.10	123	1.2.3.4
21	TR	TE	162	0.335	0.10	155	1.2.3
22	VL	VL	258	0.417	0.10	125	1.2.3.4
23	VL	TE	62	0.181	0.05	174	1.2.3.4
24	TR	TR	535	0.558	0.25	84	1.2.3.4
25	TR	TR	2931	1.479	0.90	57	2.3.4
26	TE	TE	952	0.889	0.80	99	1.2.3.4

Following similar principles, the main idea of the present paper is to study the possibility of using MLP Neural Networks to evaluate the whole concentration versus time profile at several cross-sections of a river under various flow conditions, using as little information about the river system as possible. In contrast with the earlier neural networks based work on longitudinal dispersion coefficients, this new approach relies more heavily on measurements of concentration collected during tracer tests over a range of flow conditions (for training the neural networks), but fewer hydraulic and morphological data are needed. Of course, this sort of fieldwork is expensive to carry out, so the proposed method may be of rather limited applicability in the present form. However, the work reported here should be considered as a first step, which if successful would encourage the pursuit of a similar approach based on more easily available information, such as river sinuosity and channel slope, to enable concentration versus time profiles to be predicted for different rivers, including those where neither tracer tests nor detailed channel measurements were available. This is the proposed subject of further work.

2 Experiments on the Murray Burn

Between 1999 and 2001, 26 tracer experiments (denoted E1–E26) were performed on the Murray Burn, which is a small river that flows through the Heriot-Watt University Campus at Riccarton in Edinburgh. These experiments were

undertaken at various flow rates in a 540m long reach. Each experiment consisted of the release of a tracer (Rhodamine WT) followed by the collection of tracer concentrations at up to 4 cross-sections (denoted CS1–CS4), see Fig. 1. The tracer was injected to the river at the same place each time. In the first few experiments concentration measurements were collected only at the first two cross-sections, but in the later experiments three or four of the cross-sections were used. In a few cases, equipment or human failure led to data not being successfully collected. Further information on the experiments is available in Burke (2002).

In the present paper MLP Neural Networks are used to predict the tracer concentration versus time profiles at all the cross-sections where tracer data was collected, for 19 of the experiments, which were originally considered to contain reliable data, although, as will be described in the following sections, are not necessarily free from errors. Only the selected 19 experiments are considered further in the paper, but the notation E1–E26 is retained.

The concentration measurements were taken at different sampling intervals, ranging from 20 s to 3 min, depending on the flow rate during the particular experiment and the distance of the cross-section from the tracer injection site. Also according to the flow rate, which ranged from 14 to 2931 l/s, different amounts of tracer were released to the river – varying between 0.05 and 0.9 g. Clearly, the number of samples collected during particular experiments differed, being dependent on the flow rate, the sampling interval and number of

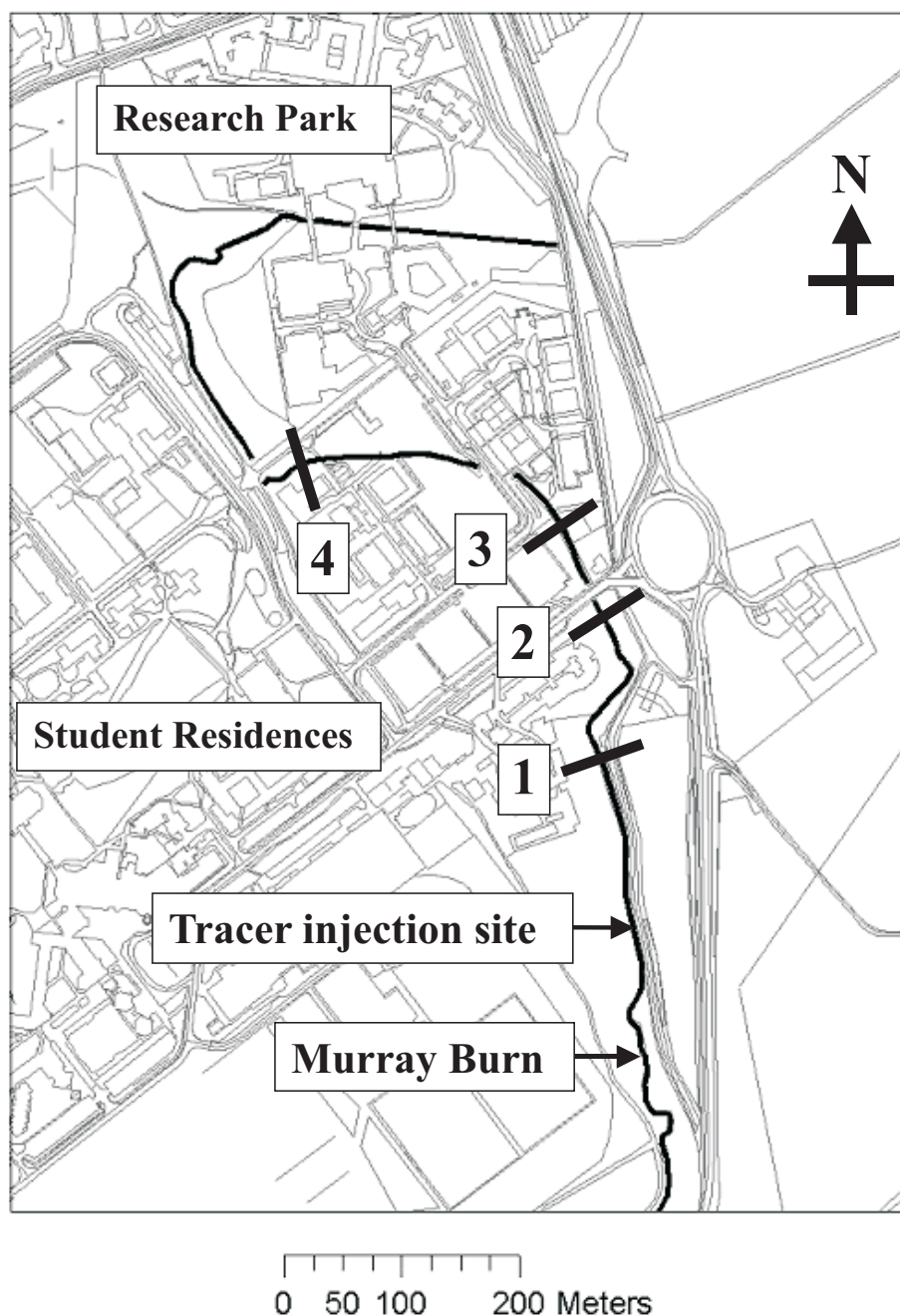


Fig. 1. Map of Murray Burn passing through Heriot-Watt University Campus: cross-sections where tracer data was collected are labelled 1–4; River flows from South to North.

cross-sections covered, however, for all experiments the sampling was designed to capture well resolved profiles. Some pertinent information is summarized in Table 1.

3 Model input variables

As stated earlier, the application of MLP Neural Networks requires several experiments to be performed to obtain a

sufficient number of data to train the model. On the other hand, the method could use a very small number of different variables. In the present study the only data used were concentration measurements collected at 4 cross-sections, distances between the cross-sections and the injection site, time, as well as flow rate and water velocity, obtained according to the data measured at the 1st and 2nd cross-sections. There are good reasons for using only a few input variables. For

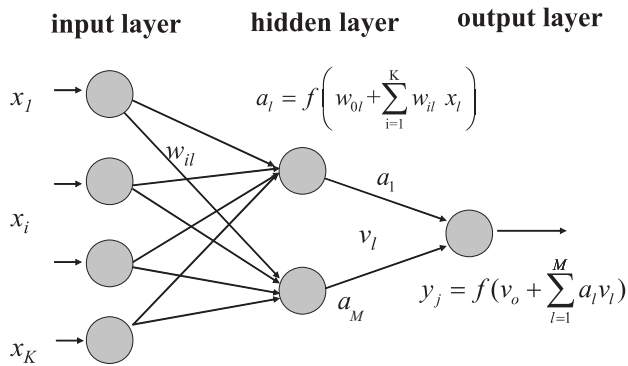


Fig. 2. Multi-Layer Perceptron Neural Network scheme.

example, using a small number of input and hidden nodes in the neural networks (see later comments) facilitates parameter optimization. There is always a contradiction between providing the model with relevant and useful information and not over-expanding the number of parameters to be optimized. Finally, it is worth noting that the basis for choosing the input variables lies on clear physical arguments, as described below. The possibility of using some automatic procedure for input selection was not thought to be appropriate.

The flow rate used was evaluated at CS1 using dilution gauging, i.e. by dividing the mass of tracer released by the area under the concentration versus time profile, and the velocity was evaluated from the centroid travel time between CS1 and CS2.

As we are interested in the evaluation of tracer versus time concentration profiles at 4 different cross-sections, their distance from the tracer release location and time are obvious input variables. Although flow rate and velocity are expected to be correlated, after preliminary tests both were considered necessary to obtain reasonable results. This reflects the fact that flow rate affects concentrations via dilution, while velocity affects the travel time of the tracer. Some complications were apparent in the flow rates and velocities evaluated at each cross-section and in each sub-reach, respectively. For example, both flow rate (obtained from dilution gauging) and velocity, from the same experiment, varied slightly from one cross-section to another. Since there was no consistent pattern to the flow rate variations it is difficult to explain them, and they were probably the result of errors due to one or more of: natural tributary inflows and/or surface water drainage from the surrounding Campus (although none were clearly visible during the experiments); non-conservative behaviour of the tracer (not considered very likely in view of the short contact times); poor mixing of the tracer; errors in the concentration profiles associated with inconsistent background

Table 2. Neural Network's input and output variables. DST – Distance from injection point; VEL – water VELOCITY collected between CS1 and CS2; FLW – FLOW collected at CS1; TBP – Time Before the Peak – in min; TAP – Time After the Peak – in min; CRL – Concentration in Rising Limb as a % of peak concentration; CFL – Concentration in Falling Limb as a % of peak concentration; PTT – Peak Travel Time; PCC – Peak ConCetration.

Network	N1	N2	N3	N4
Input variables	DST	DST	DST	DST
	VEL	VEL	VEL	VEL
	FLW	FLW	FLW	FLW
	TBP	TAP		
Output variables	CRL	CFL	PTT	PCC

concentrations, incomplete profiles and/or poorly resolved profiles. Variations in velocity along the experimental reach could also arise from these errors, but would also reflect the non-uniform nature of the river channel – generally the slope reduces and the channel narrows and deepens with distance from the tracer release site.

Although the evaluated flow rates and velocities showed an expected positive correlation over the full range of experiments (velocity increasing non-linearly with flow rate in the first sub-reach, for example), some inconsistencies were also apparent between experiments conducted at similar flow rates. For example, consider the flow rates at CS1 and the velocities in the first sub-reach for the following four pairs of experiments shown in Table 1: E10 and E15; E16 and E17; E9 and E18; E20 and E22; and four experiments performed in almost equal flow conditions: E4, E5, E8 and E15. We would expect little variation in velocity if there is little variation in flow and we would expect higher velocities to be associated with higher flow rates. The fact that this isn't always the case reflects the presence of noise in the data that will tend to decrease the quality of the predictions being made. To avoid using the mass of tracer injected to the river as an input variable (potentially necessary because different masses of tracer were used in the experiments, based on a visual assessment of the flow rate before each experiment), all the concentration data were normalized to values that corresponded to the injection of 0.1 g of tracer. The predicted results, however, are presented (re-scaled) for the real injected tracer mass. This approach was possible due to the assumption of a linear dependence between the mass of tracer injected and the concentration of solute in the river. Because the channel size and shape were essentially the same for all the experiments, no additional input variables catering for channel geometry were needed.

Note that Table 1 also shows two divisions of the data into training, validation and testing sets. The training data are used by the neural networks to optimize the model parameters, the validation set is used by the stopping criterion of the

Table 3. Neural Network's technical details.

Data division	MLP network	Number of parameters	Network structure
D9	N1	37	4–6–1
D9	N2	37	4–6–1
D9	N3	16	3–3–1
D9	N4	16	3–3–1
D5	N1	31	4–5–1
D5	N2	31	4–5–1
D5	N3	16	3–3–1
D5	N4	16	3–3–1

optimization algorithm (to avoid overfitting of the network), and the testing set includes fully independent data, used to check the quality of predictions made with the model. In the present work all of the data collected during a particular experiment belonged to the same data division set. Two versions of data division were used – the first one (D9) with 9 experiments in the training, 5 in the validation and 5 in the testing set allowed more data for optimization procedures, in principle enabling better prediction, whereas the second one (D5) limited the training set to 5 experiments and the validation set to only 3 experiments, leaving most of the data (11 experiments) in the testing set. Further reductions in the amount of training data did not produce reliable results.

4 Methodology of proposed approach

In some preliminary studies we found that the prediction of whole concentration versus time profiles at several locations downstream from the injection point by means of only one neural network did not produce accurate predictions. It turned out to be better to use different optimization criteria for the evaluation of the peak time, the peak tracer concentration and concentrations at other times. Also it was found that different neural network structures should be used to model each of the features of concentration profiles mentioned above. There is also evidence in the literature (Wagener et al., 2002; Wörman et al., 2005; Wörman and Wachniew, 2007) that different parts of the concentration versus time profile are governed by different physical processes, which are generally modelled by separate terms in the case of physically-based approaches, such as the Transient Storage Model. In this regard, for example, one may distinguish the rising limb, the peak region, the falling limb and the far-limb of the concentration versus time profile. Therefore, it is likely that a very promising approach is to split the problem into a set of sub-problems. Hence, instead of using one single neural network to predict all the features of a concentration versus time profile, it is proposed that separate networks be developed to predict specific features of the profile. In particular, the four most useful features of concentration versus

Table 4. Equations of power regression of flow (Q), peak travel time (PTT) and unit peak concentration (CUP), obtained for each cross-section separately. The subscripts in PTT, Q and CUP refer to the particular cross-section.

Data selection	D5	D9
Time to peak	$PTT_1=208.94Q_1^{-0.66}$	$PTT_1=226.06Q_1^{-0.68}$
	$PTT_2=258.94Q_2^{-0.55}$	$PTT_2=260.61Q_2^{-0.56}$
	$PTT_3=362.57Q_3^{-0.55}$	$PTT_3=379.47Q_3^{-0.57}$
	$PTT_4=269.16Q_4^{-0.43}$	$PTT_4=269.16Q_4^{-0.43}$
Unit peak concentration	$CUP_1=3.11PTT_1^{-1.06}$	$CUP_1=2.93PTT_1^{-1.01}$
	$CUP_2=3.28PTT_2^{-0.95}$	$CUP_2=3.73PTT_2^{-0.99}$
	$CUP_3=5.02PTT_3^{-1.03}$	$CUP_3=4.96PTT_3^{-1.02}$
	$CUP_4=3.47PTT_4^{-0.81}$	$CUP_4=3.47PTT_4^{-0.81}$

time profiles at a particular cross-section, which should be available from the model, are the peak concentration, the arrival time of the peak and the shapes of the rising and falling limbs of the profile. Note that the far-limb, which, if present, is characterised by elevated concentrations in the later part of the profile tails, was not observed in the Murray Burn experiments. So it was not necessary to include it in the present study, but it may turn out to be needed when the study is expanded to include data from different rivers. Following this idea, in the proposed approach four separate networks are used: N1 concerns the profile's rising limb (i.e. concentrations before the peak); N2 concerns the profile's falling limb (i.e. concentrations after the peak); N3 concerns the peak travel time; and N4 concerns the peak concentration. Two model versions were constructed: version V1 consisted of networks N1 – N4; version V2 consisted of networks N1 – N3. The peak concentration in V2 was evaluated in the way described below. The input and output variables of the four neural networks are summarized in Table 2.

It is important to note that for networks N1 and N2 no information about the peak concentration is needed. The output of the network is given as a percentage of the peak concentration, with the actual concentrations being obtained when the peak concentration is evaluated.

In version V1, after merging the results evaluated by networks N1 – N4, one obtains the full concentration versus time profile. Unfortunately, as each step is performed separately by a different network, the V1 approach does not necessarily conserve the mass of tracer injected. This could be the result of either the peak concentration or the limbs of the profile not being accurately predicted. The former is avoided in version V2 because no neural network is used to evaluate the peak concentration. Instead, when the shapes of the concentration versus time profile are obtained from networks N1 and N2, the peak concentration is chosen such that the areas under the curves are equal at all cross-sections. Assuming that the flow is the same along the whole experimental reach (only flow from CS1 is used by the model), this

Table 5. Comparison of peak concentration (PCC) and travel time (PTT) obtained from V1, V2 and UPA methods: highlighted values indicate best of UPA, V1 and V2 for each cross-section, and for each estimated parameter/error criterion, separately. ME – Mean of absolute Error; RMSE – Root Mean Square Error.

Case	Cross-section	PCC RMSE			PCC ME			PTT RMSE		PTT ME	
		UPA	V1	V2	UPA	V1	V2	UPA	V1-2	UPA	V1-2
D5	CS1	0.373	0.219	0.952	0.040	0.030	0.034	1.673	1.967	0.043	0.058
	CS2	0.532	0.408	0.167	0.062	0.032	0.029	3.471	2.479	0.046	0.041
	CS3	0.265	0.342	0.155	0.061	0.033	0.032	4.637	4.176	0.047	0.043
	CS4	—	0.256	0.143	—	0.046	0.038	—	6.175	—	0.024
	CS4*	0.791	0.303	0.094	0.092	0.055	0.019	10.816	4.944	0.049	0.027
D9	CS1	0.407	0.633	0.674	0.031	0.027	0.026	1.559	1.678	0.043	0.048
	CS2	0.457	0.290	0.264	0.055	0.028	0.029	3.457	1.658	0.044	0.032
	CS3	0.228	0.175	0.165	0.055	0.022	0.025	3.539	2.934	0.043	0.030
	CS4	—	0.356	0.073	—	0.067	0.021	—	3.372	—	0.026
	CS4*	0.791	0.587	0.095	0.092	0.088	0.012	10.816	4.131	0.049	0.029

approach provides mass conservation and makes the modeling easier, because one network is excluded. Note, however, as already alluded to, that unexplained errors in some of the observed concentration profiles led to real or apparent non-conservative behaviour in some of the experiments. This has some impact on the quality of the agreement between the predicted and observed profiles, as discussed in Sect. 6.

5 Multi-layer perceptron neural networks

MLP Neural Networks are well known universal approximators Haykin (1999). They consist of nodes grouped in input, hidden and output layers. Nodes in consecutive layers are connected via weights, which are the parameters to be optimized. A typical MLP network scheme is shown in Fig. 2 in which x_i represent the K input variables, y_j represents the output variable, w_{il} and v_l are network weights to be optimized and a_l represent the signals dispatched by each of M hidden nodes.

In the present application the signals dispatched by each of the hidden nodes (a_l) are evaluated via the following so-called logistic function:

$$a_l = \left[1 + \exp \left(-w_{0l} - \sum_{i=1}^K w_{il} x_i \right) \right]^{-1} \quad (1)$$

In the final node, the output from the network is evaluated via a linear function (in case of N3 and N4),

$$y_j = -v_0 - \sum_{l=1}^M v_l a_l \quad (2)$$

or via a logistic function (for N1 and N2)

$$y_j = \left[1 + \exp \left(-v_0 - \sum_{l=1}^M v_l a_l \right) \right]^{-1} \quad (3)$$

When applying the logistic function to the final node, the output value is always constrained to be within a limited range. This is favourable when predicting the shape of the profile, because the concentration in the rising or falling limb of the profile cannot exceed the peak concentration or drop below 0. However, this is not suitable for evaluating the peak travel time or the peak concentration, so the linear function is used in the output node in these cases.

In each case all the available input and output data were normalized linearly to [0,1] before being presented to the network. The computed value of the output variable was then re-scaled, to obtain the final output value. The number of hidden nodes was chosen experimentally. For the evaluation of peak travel time (N3) and peak concentration (N4), the number of training data used was significantly smaller than in case of modelling the shape of the profile (N1 and N2), because there was only one piece of data per cross-section and per experiment. As a result, to not overparameterize the network, it was assumed that no more than 3 hidden nodes should be used. In case of profile shape modelling (N1 and N2) the number of data available at a particular cross-section during each experiment differed according to the velocity, the spread of the concentration profile and the time interval between consecutive concentration values, but the number of pieces of data was high enough to not impose a priori any limit on the number of hidden nodes to be used. In practice the best results were obtained for neural networks with 5–6 hidden nodes.

For the optimization process the following well known objective function (J) was applied:

$$J = \min_{\mathbf{v}, \mathbf{w}} \sum_{j=1}^N (d_j - y_j(\mathbf{w}, \mathbf{v}))^2 \quad (4)$$

where d_j is the measured value of the output variable, y_j is the predicted value of the output variable and N represents the number of data. To optimize each network the Levenberg-Marquardt algorithm (Press et al., 1990) with a multi-start approach (to avoid sticking in local optima) was applied. For each start of the algorithm, the initial parameter values were chosen randomly.

Table 3 presents the details of the MLP Neural Networks used to evaluate tracer concentration versus time profiles for the two data divisions, D5 and D9. The network structure shown presents the number of input, hidden and output nodes, consecutively.

6 Results

Predictions of concentration versus time profiles were made using both versions (V1 and V2) of the proposed approach. Results were obtained for all 4 cross-sections of the 19 experiments summarized in Table 1, using the two data divisions described earlier (D5 and D9).

Obviously in the case of peak travel time and peak concentration, only one measurement per experiment and cross-section exists. This may cause problems when comparing results, because the most popular comparison index, the root mean square error (RMSE), also used in the present paper, is more sensitive to larger magnitude variables. For example, it would take exaggerated values when considering concentrations at CS1 compared to CS4 (because concentrations decrease with distance from the tracer release site) and it would take exaggerated values when considering concentrations at lower flows compared to those at higher flows (because of smaller dilution). Similar arguments can be made for travel times also. Hence, another comparison index, namely Mean of Absolute Error (ME), in the form proposed by Kashefipour and Falconer (2002) for longitudinal dispersion coefficient estimation, a topic in which similar problems occur, is also used. ME is defined as:

$$ME = \frac{1}{N} \sum_{j=1}^N \left| \log_{10} \left(\frac{y_j(\mathbf{w}, \mathbf{v})}{d_j} \right) \right| \quad (5)$$

where N is the number of data, d_j and y_j are measured and predicted values, respectively, and w and v are as previously defined in Sect. 5.

When comparing results from networks N1 and N2, which predict the rising and falling limbs of the profile, the ME criterion is less important because plenty of data were collected during each experiment and the measurements are always in

the [0,1] interval, where 1 indicates 100% of peak concentration. As a result, only the RMSE criterion was considered for these cases. Although both RMSE and ME are useful in general evaluations of different methods, visual comparisons remain an effective way for comparing predicted and observed concentration versus time profiles.

In order to assess how well the neural network approach performs, it is useful to compare it against an existing method. The method suggested by Wallis (2005), designed to evaluate peak travel time and peak concentration for each cross-section separately as a function of flow rate, was available for this purpose. Like the proposed neural network approach, it requires several tracer tests to be performed under different flow rates. For each cross-section, a nonlinear relationship between flow and peak travel time is established by means of non-linear regression and then the unit peak attenuation curve concept (Jobson, 1997; Wallis, 2005) is used. From the unit peak concentration one simply finds the true peak concentration value by re-scaling. This method is referred to as the UPA method in the remainder of the paper.

In the application of the UPA method, the required non-linear trends were identified for each cross-section for both variants of data division (D5 and D9), using only the data in the neural network training set. A power law type of regression was selected, for both peak travel time versus flow and for unit peak concentration versus peak travel time. These regression equations are shown in Table 4.

In contrast to the neural network approach in which only flow rate at CS1 was used, individual flow rates at each cross-section (evaluated by dilution gauging) were used in the UPA regression analyses. Unfortunately, for experiments E16 and E17 difficulties with the tails of the profiles at CS4 meant that the flows were unreliable. As a result of this and also because data was not recorded at CS4 in all the experiments, the training data set that could be used in the UPA method at the fourth cross-section was the same for both D5 and D9 data divisions. This reduced data set, denoted as CS4* was comprised of experiments E20, E24 and E25, see Table 1. Hence the regression equations are the same for both data divisions, see Table 4.

Results are presented in Tables 5–7 and in Figs. 3 and 4. Table 5 shows results of RMSE and ME for peak concentration (PCC) for each cross-section from three methods (V1, V2 and UPA) and for both data divisions (D5 and D9). Values of RMSE and ME were evaluated from all the cases in the training, validation and testing data sets. Similarly, Table 5 also shows results for peak travel time (PTT) for which results from V1 and V2 are the same (denoted as V1-2). Table 6 shows the neural network results of RMSE and ME for peak concentration and peak travel time in a different way – results are evaluated over all cross-sections for each of the training, validation and testing data sets. Table 7 shows RMSE values for the rising and falling limbs computed from neural networks N1 and N2, respectively, for both data divisions: again results are evaluated over all cross-sections for

Table 6. Comparison of peak concentration (PCC) and peak travel time (PTT) obtained from V1 and V2 methods: highlighted values indicate better of V1 and V2 for each data division, and for each error criterion/data set, separately. ME – Mean of absolute Error; RMSE – Root Mean Square Error; TR – TRaining data set; VL – VaLidation data set; TE – TEsting data set.

Prediction for	Case	Tool	ME TR	ME VL	ME TE	RMSE TR	RMSE VL	RMSE TE
PCC	D5	V1	0.027	0.033	0.038	0.175	0.168	0.400
		V2	0.024	0.021	0.040	0.161	0.090	0.704
	D9	V1	0.027	0.017	0.060	0.199	0.071	0.840
		V2	0.021	0.022	0.039	0.128	0.093	0.796
PTT	D5	V1-2	0.019	0.015	0.063	1.020	0.687	4.766
	D9	V1-2	0.038	0.023	0.044	2.908	0.996	3.841

Table 7. Root Mean Square Error of the rising and falling limbs of the concentration versus time profile: highlighted values indicate better of D5 and D9 for each data set. TBP – Time Before the Peak – in minutes; TAP – Time After the Peak – in minutes; TR – TRaining data set; VL – VaLidation data set; TE – TEsting data set.

Shape	TR	VL	TE
TBP D5	0.0727	0.1215	0.1043
TBP D9	0.0705	0.0978	0.0983
TAP D5	0.0441	0.0430	0.0510
TAP D9	0.0396	0.0373	0.0500

each of the training, validation and testing data sets. To aid comparisons in the tables, the best result (minimum RMSE and ME) is highlighted. Figs. 3 and 4 show concentration profiles for all the experiments and allow predictions from neural network methods V1 and V2 to be compared with the observed profiles.

Table 5 shows that, in general, methods V1 and V2 make better predictions of peak concentration and peak travel time than the UPA method, for both data divisions. Only for CS1 does the UPA method sometimes give better results. Predictions of peak concentration from method V2 are generally better than those from method V1. In the majority of cases using either RMSE or ME identifies the same best method. The peak travel time results illustrate well the reduced sensitivity of ME, compared to that of RMSE, to the magnitude of the variable under scrutiny. Here, if this sensitivity were not recognised the RMSE results would suggest a significant deterioration in the quality of the peak travel time predictions from both the UPA and neural network approaches with distance along the reach. Yet this is not supported by Figs. 3 and 4. In contrast, the ME results better reflect the quality of the peak travel time predictions. Indeed, it is not even clear that there is any trend in the quality of the predictions of peak travel time or peak concentration with distance along the reach. It is also important to remember that many fewer

tracer tests covered CS3 and CS4 than CS1 and CS2, so the results for far lying cross-sections may be more sensitive to measurement errors in the concentration profiles.

In general the results in Table 5 show some improvement in predictions using data division D9 compared to D5. In about 65% of the cases there are reduced values of ME and RMSE, but there is a large variation in the magnitude of the improvement and there are several inconsistencies also. This indicates that the quality and consistency of the data used as well as the number of data used has a bearing on the results. For example, moving an experiment containing relatively poor data into the training data set of a neural network may have an adverse effect on the predictions. Similarly, introducing an outlying piece of data into the UPA method may result in less representative regression relationships being identified that go on to have a negative effect on predictions.

Table 6 shows that the neural network predictions of peak concentration and peak travel time are generally poorer for the testing data set than for the training and validation sets. This trend is found for both versions (V1 and V2), for both performance criteria (ME and RMSE) and for both data divisions (D5 and D9). Of course, this is to be expected since most models perform better with the data used in their calibration than when using independent data. As with the results in Table 5 using either ME or RMSE identifies the same better model, although Table 6 does not show so clear a trend that network V2 gives better predictions than network V1. There are no clear differences between the two data divisions.

When comparing RMSE of the rising and falling limbs of the concentration profiles, see Table 7, a small but consistent improvement is evident when more data are used for training, i.e. results are always better for data division D9 than for D5. As in Table 6, results are generally poorer for the testing data than the calibration data.

Figures 3 and 4 show greater detail of the quality of the predictions made by V1 and V2 and show that in the majority of cases both networks are successful in reproducing the main features of the profiles. The figures provide the

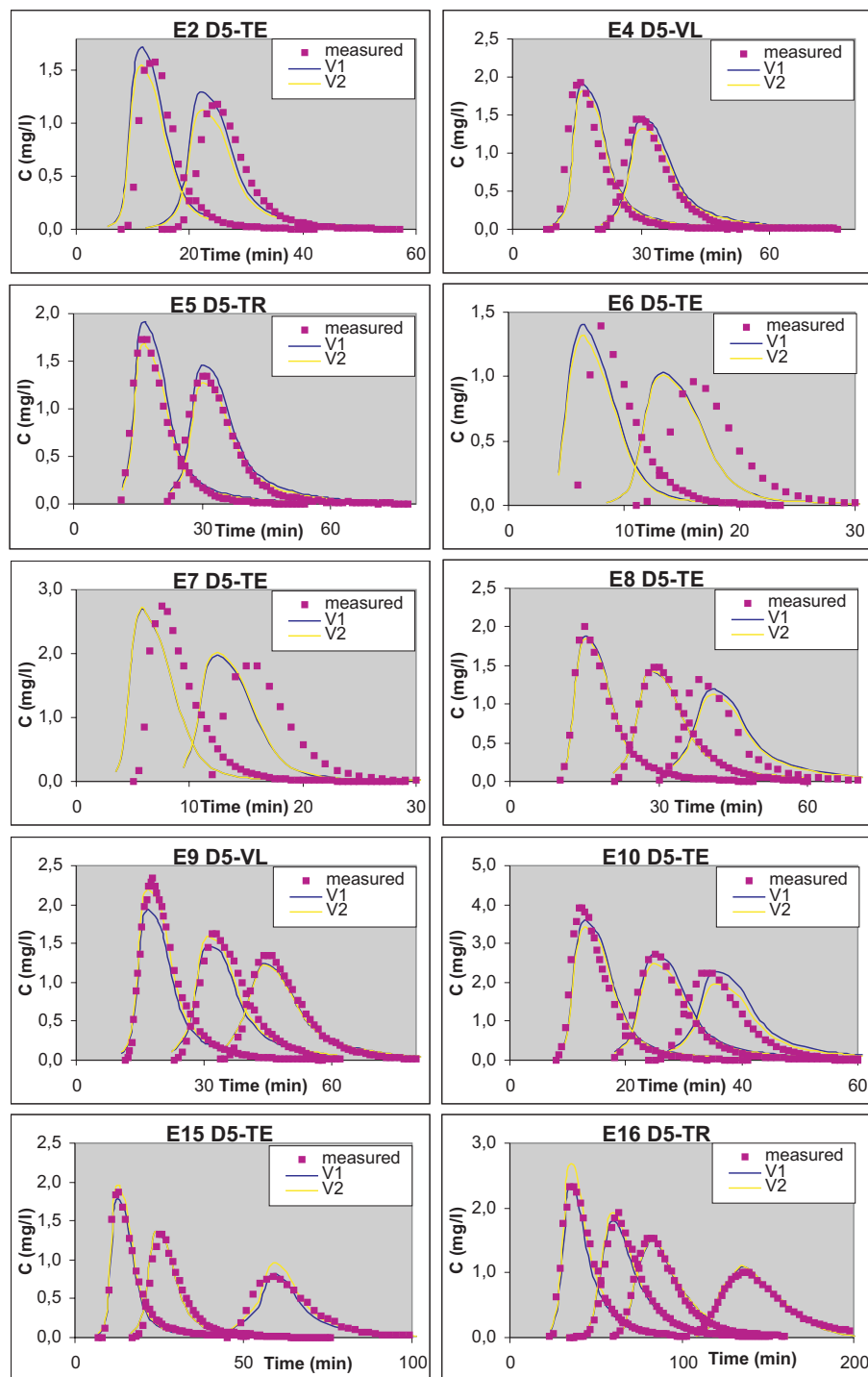


Fig. 3a. Concentration versus time profiles for data division D5, experiments E2–E16. The title of each plot gives the experiment number, the data division case and the data set type in which the experiment was included.

detail on which the earlier comments on the performance of the neural networks were based and therefore support those ideas. For example, predictions are generally poorer using the testing data than the calibration data, and there are some improvements when more data is used in the training data set

(compare Fig. 3 with Fig. 4). It is also clear that the weak link in the predictions is the peak travel time, indicated by the frequent phase errors, whereas the peak concentrations and profile shapes are generally well predicted.

Some of the poorer predictions are associated with the

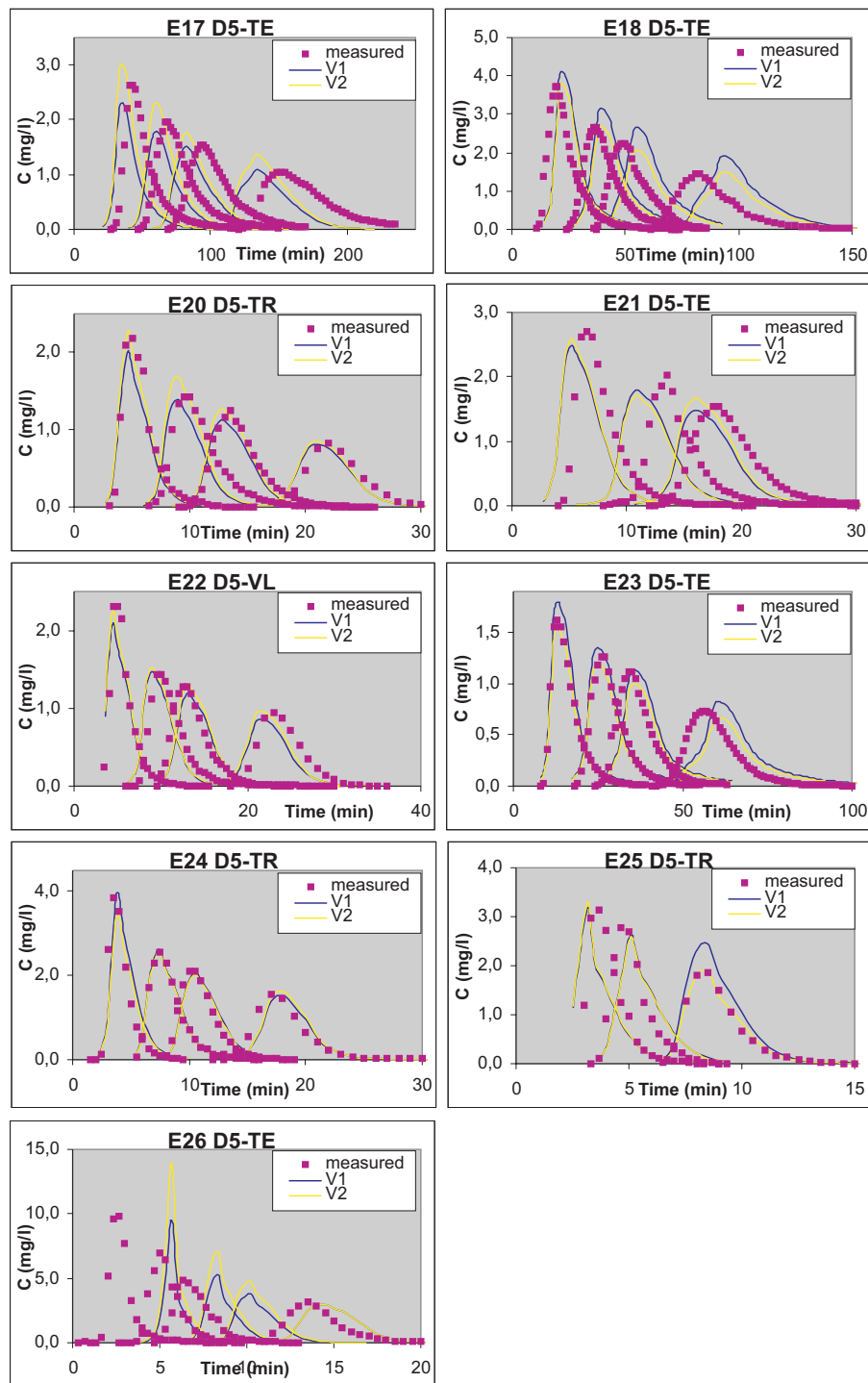


Fig. 3b. Concentration versus time profiles for data division D5, experiments E17–E26. The title of each plot gives the experiment number, the data division case and the data set type in which the experiment was included.

highest flow cases, namely experiments 24, 25 and 26. This is probably a reflection of several things. Firstly, the concentration profiles were collected at a poorer temporal resolution than the others, so are not very well defined. Sec-

ondly, although two of these three cases were in the training data sets for both data divisions, there were few high flow cases to learn from. This should also be treated as a warning not to extrapolate the results far beyond the calibration

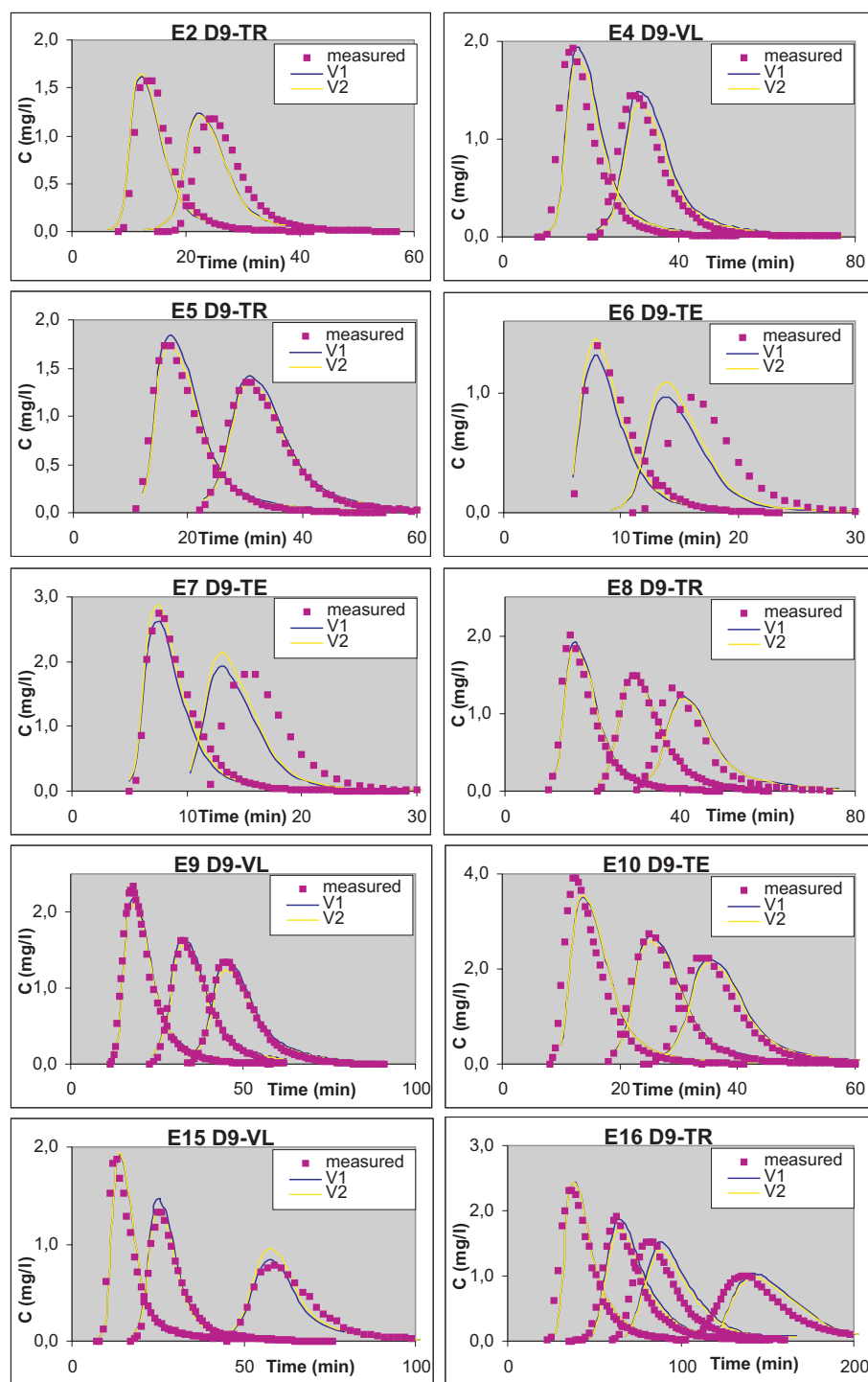


Fig. 4a. Concentration versus time profiles for data division D9, experiments E2–E16. The title of each plot gives the experiment number, the data division case and the data set type in which the experiment was included.

data range, which is a well known problem in environmental modelling, particularly when using non-physically based models. Thirdly, predicting events under extreme conditions

is always likely to be more difficult than predicting events closer to mean conditions, which tend to be better covered by the majority of data collected. It is interesting that the

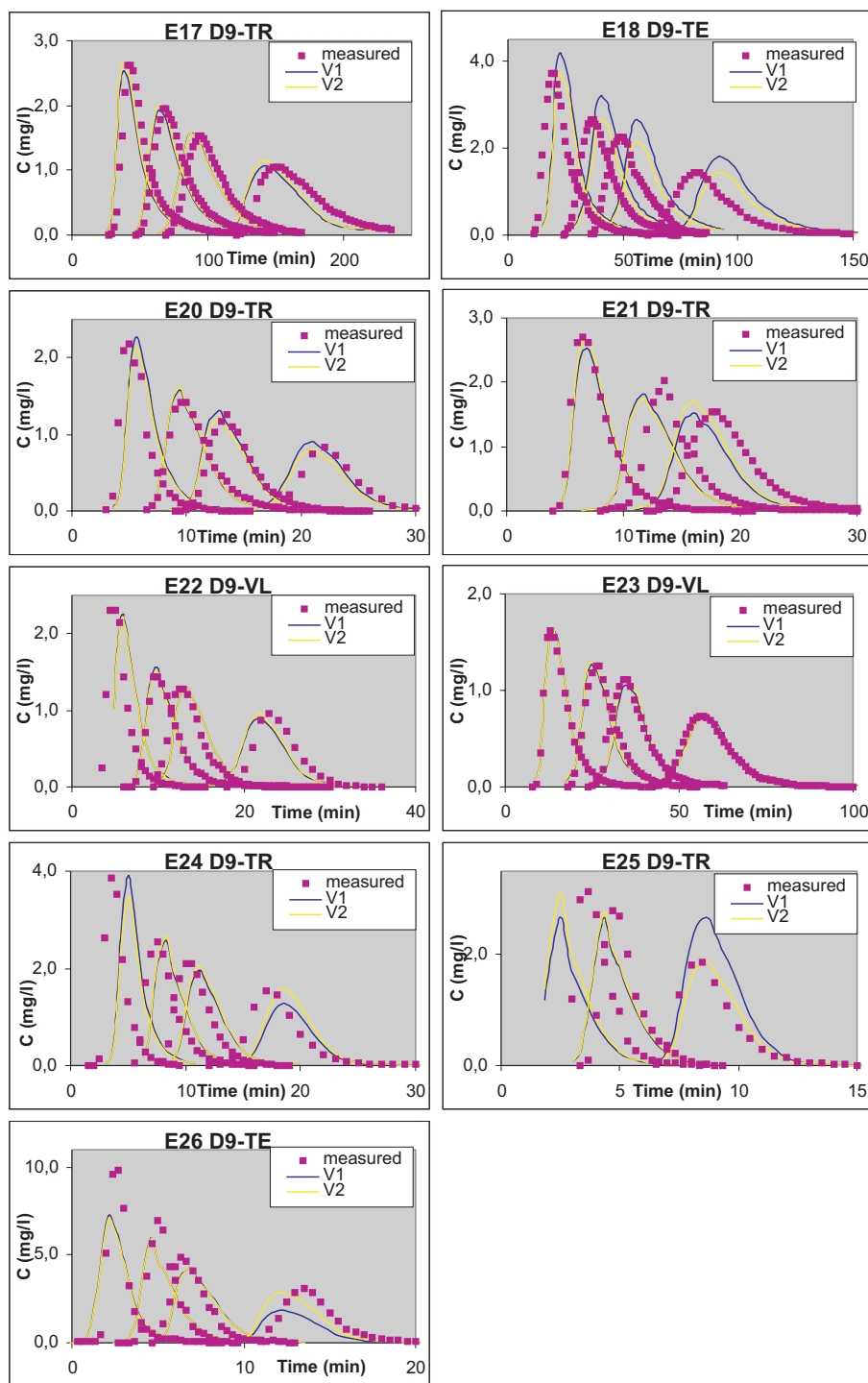


Fig. 4b. Concentration versus time profiles for data division D9, experiments E17–E26. The title of each plot gives the experiment number, the data division case and the data set type in which the experiment was included.

prediction for experiment 26 is considerably improved using more experiments in the training data set. In contrast, however, predictions for experiments 24 and 25 are better when

there are fewer experiments in the training data set.

Finally, it is worth noting that it is the occasional very large prediction error, especially in peak concentration at CS1 in

E26, that can have a significant effect on the RMSE error criteria (compare results in Table 5 for V1 and V2 for both data divisions).

7 Conclusions

In this paper Multi-Layer Perceptron Neural Networks were applied to the evaluation of solute concentration versus time profiles in a small river. The data were collected at four cross-sections of the Murray Burn, located in Edinburgh, during several tracer experiments performed under different hydrological conditions. The proposed approach used separate neural networks for the evaluation of peak concentration, peak travel time and both the rising and falling limbs of the concentration profiles. The only input data needed were the concentration profiles, estimates of flow and water velocity (obtained from the concentration data collected at the 1st and 2nd cross-sections), time and the distance of the cross-sections from the release site.

In general the neural networks predicted the main features of the concentration profiles satisfactorily, although the prediction of the peak travel time was disappointing in several cases. The predicted peak concentrations were generally better than those obtained using a unit peak attenuation method, and a method with mass-conservation assured generally performed better than a method that did not account for mass-conservation. Predictions of peak travel time were also better using the neural networks than the unit peak attenuation method. Including more data into the neural network training set clearly improved the prediction of the shapes of the concentration profiles. Similar improvements in peak concentration were less significant and the travel time prediction appeared to be largely unaffected.

Acknowledgements. Funding for this research was provided by Polish Ministry of Science and Higher Education, Grants no. 2 P04D 026 29 and 2P04D 009 29. The second author is grateful to the Royal Society of Edinburgh, the Polish Academy of Sciences and the British Embassy in Warsaw for funding visits during which work contributing to the paper was undertaken.

Edited by: A. Montanari

References

- Burke, N. J.: Travel time and flow characteristics of a small stream system, PhD Thesis, Heriot-Watt University, Edinburgh (unpublished), 2002.
- Cheong, T. S. and Seo, I. W.: Parameter estimation of the transient storage model by a routing method for river mixing processes, *Water Resour. Res.*, 39(4), 1074–1084, 2003.
- Czernuszenko, W. and Rowiński, P. M.: Properties of the dead-zone model of longitudinal dispersion in rivers, *J. Hydraul. Res.*, 35(4), 491–504, 1997.
- Czernuszenko, W., Rowiński, P. M., and Sukhodolov, A.: Experimental and numerical validation of the dead-zone model for longitudinal dispersion in rivers, *J. Hydraul. Res.*, 36(2), 269–280, 1998.
- De Smedt, F., Brevis, W., and Debels, P.: Analytical solution for solute transport resulting from instantaneous injection in streams with transient storage, *J. Hydrol.*, 315, 25–39, 2005.
- Fischer, H. B., List, E. J., Koh, R. C. Y., Imberger, J., and Brooks, N. H.: *Mixing in Inland and Coastal Waters*, Academic Press, London, UK, 1979.
- Guymet, I. and Dutton, R.: Application of a transient storage model to meandering channel studies of solute transport and dispersion, in: *Water Quality Hazards and Dispersion of Pollutants*, edited by: Czernuszenko, W. and Rowiński, P. M., Springer, New York, USA, 85–107, 2005.
- Hart, D. R.: Parameter estimation and stochastic interpretation of the transient storage model for solute transport in streams, *Water Resour. Res.*, 31(2), 323–328, 1995.
- Haykin, S.: *Neural Networks: A Comprehensive Foundation*, Macmillan College Publishing Co., New York, USA, 1999.
- Jobson, H. E.: Predicting travel time and dispersion in rivers and streams, *J. Hydraul. Eng.-ASCE*, 123(11), 971–978, 1997.
- Kashefipour, S. M., Falconer, R. A., and Lin, B.: Modelling longitudinal dispersion in natural flows using ANNs, in: *Proceedings of River Flow 2002*, Louvain-la-Neuve, Belgium, 111–116, 2002.
- Kashefipour, S. M. and Falconer, R. A.: Longitudinal dispersion coefficients in natural channels, *Water Res.*, 36(7), 1596–1608, 2002.
- Manson, J. R.: Predicting stream transient storage zone characteristics from tracer data, *Water Resour. Res.*, 36(2), 629–631, 2000.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T.: *Numerical Recipes in C, The Art of Scientific Computing*, Cambridge University Press, Cambridge, UK, 1990.
- Rowiński, P. M., Napiórkowski, J. J., and Owczarczyk, A.: Transport of passive admixture in a multi-channel river system - the Upper Narew case study, Part 2, Application of dye tracer method, *Ecology and Hydrobiology*, 3(4), 381–388, 2003.
- Rowiński, P. M., Piotrowski, A., and Napiórkowski, J. J.: Are artificial neural network techniques relevant for the estimation of longitudinal dispersion coefficient in rivers?, *Hydrolog. Sci. J.*, 50(1), 175–187, 2005.
- Rutherford, J. C.: *River Mixing*, Wiley, Chichester, UK, 1994.
- Taylor, G. I.: The dispersion of matter in turbulent flow through a pipe, *Proceedings of the Royal Society of London*, A223, 446–468, 1954.
- Thackston, E. L. and Schnelle, K. B. J.: Predicting effects of dead zones on stream mixing, *Journal of the Sanitary Engineering Division, Proceedings of the American Society of Civil Engineering*, 96(SA2), 319–331, 1970.
- Wallis, S. G.: Experimental study of travel times in a small stream, in: *Water Quality Hazards and Dispersion of Pollutants*, edited by: Czernuszenko, W. and Rowiński, P. M., Springer, New York, USA, 109–120, 2005.
- Wallis, S. G., Piotrowski, A., Rowiński, P. M. and Napiórkowski, J. J.: Prediction of dispersion coefficients in a small stream using Artificial Neural Networks, *Proceedings of the 32nd IAHR Congress, Venice*, paper B2b-083-O, 2007.
- Wagener, T., Camacho, L. A. and Wheeler, H. S.: Dynamic identifiability analysis of the transient storage model for solute transport

- in rivers, *J. Hydroinform.*, 4(3), 199–211, 2002.
- Wörman, A., Klove, B., Wachniew, P., Czuprynski, P. and Packman, A.: Tracer test in Hobol creek, Norway, under different flow conditions, *Acta Geophysica Polonica*, 53(4), 517–526, 2005.
- Wörman, A. and Wachniew, P.: Reach scale and evaluation methods as limitations for transient storage properties in streams and rivers, *Water Resour. Res.*, 43, W10405, doi:10.1029/2006WR005808, 2007.